

Smart Detector Cell: A Scalable All-Spin Circuit for Low Power Non-Boolean Pattern Recognition

Hamidreza Aghasi[†], Rouhollah Mousavi Iraei^{*}, Azad Naeemi^{*} and Ehsan Afshari[†]

Abstract

Abstract—We present a new circuit for non-Boolean recognition of binary images. Employing all-spin logic (ASL) devices, we design logic comparators and non-Boolean decision blocks for compact and efficient computation. By manipulation of fan-in number in different stages of the circuit, the structure can be extended for larger training sets or larger images. Operating based on the mainly similarity idea, the system is capable of constructing a mean image and compare it with a separate input image within a short decision time. Taking advantage of the non-volatility of ASL devices, the proposed circuit is capable of hybrid memory/logic operation. Compared with existing CMOS pattern recognition circuits, this work achieves a smaller footprint, lower power consumption, faster decision time and a lower operational voltage.

1 Introduction

Pattern recognition and in particular, image recognition techniques have been widely studied in machine learning and image processing [1–3]. Hardware demonstration of computation units for pattern recognition; however, has consistently been a challenging problem in terms of chip size, power consumption, computation complexity and decision speed.

Among different solid state technologies, CMOS provides the chance of low cost, highly-integrated low power implementation for pattern recognition [4–6] and processing [8, 9] systems. For boolean logic systems, CMOS gates exhibit processing speeds up to a few GHz and can be designed to have a low static power. However, the dynamic power consumption of a large system with a GHz clock frequency can still limit the scalability. Fan-in and fan-out considerations for CMOS devices also impact the speed, power consumption and the size of devices. Besides boolean systems, some novel non-Boolean techniques have been developed to overcome these issues. In non-Boolean systems, logic gates will no longer be the key block and analog/mixed signal circuits are used. In [6], the authors propose a technique for non-Boolean training and detection of image pixels using a network of coupled oscillators. This structure has the capability to detect any scaled or rotated version of a desired image. On the other hand, this method suffers from high computational complexity, large area and high power consumption which limit the application for large image arrays.

[†]School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14850

^{*}School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332

To this, we should also add the long convergence time. Other proposed CMOS systems have demonstrated artificial neural networks (ANN) by designing circuits emulating neurons and synapses [4, 5]. In these systems, the larger computation burden, leaves the search open for new solutions.

To overcome the limitations of CMOS devices, other technologies are being investigated for pattern recognition applications. Spintronic devices, in particular, have received a lot of attention recently because of some unique properties, e.g., low voltage operation and non-volatility. In [10], a non-volatile logic-in-memory full adder is fabricated using the magnetic tunnel junctions (MTJ's). The proposed architecture is compared with an $0.18\mu\text{m}$ CMOS process counterpart and exhibits major advantages. The dynamic power consumption compared to a conventional CMOS circuit is 23% reduced due to reduction in the number of paths from V_{DD} to GND. On the other hand the static power consumption is eliminated due to the non-volatility and the chip area is also reduced. As stated in [11], the ASL devices are also non-volatile and the computational state is preserved when the power to the circuit is turned off. In [12], a spin-based artificial neural network (ANN) is proposed using lateral spin valves to achieve a low power consumption and a low operational voltage. In [13], spin switches to develop compact neurons and synapses are proposed. In [14], all-spin logic (ASL) and charge-spin logic (CSL) devices are shown to be capable of Boolean and non-Boolean operations which make them an attractive choice to build some fundamental blocks such as ring oscillators. Recently, the design of ASL gates with graphene channels have been proposed [15]. Due to the unique features of graphene in terms of the spin transport, the design of Boolean and non-Boolean computation units with these new devices can be investigated as a future direction. Moreover, majority gate operation of ASL devices has been previously introduced in some Boolean logic systems [19, 20]. This unique feature of these devices can overcome fan-in and fan-out limitations of large integrated systems. Besides, the inverting and non-inverting operation modes of ASL devices can be the key to design many logic circuits e.g., full adder circuits and multipliers [22]. The time domain transient behavior of magnetization in these devices can also provide another degree of freedom to demonstrate non-Boolean operations as it will be discussed later in this paper. These features, enable us to design an all-spin logic non-Boolean compact structure with low power consumption and low computational complexity.

In this paper we propose a novel pattern recognition circuit that takes advantage of most novel features of spintronic devices such as non-volatility, efficient implementation of majority gates and XOR functions and the ability to distinguish strong and weak majorities. Non-volatility of the devices enables storing large sets of training images within the logic with no standby power dissipation. This feature also enables instant-on operation and saves on energy and delay penalties imposed by loading training images from a main memory.

The rest of this paper is organized as follows. Section II describes the operation of ASL devices. The proposed approach and the basic of computation are given in Section III. In Section IV, the proposed architecture and a comprehensive discussion on design

considerations are presented. Simulation results and summary are shown in Sections V and VI, respectively.

2 All-Spin Logic Devices

Spin of electron is introduced as a new state variable in spintronic devices to process and store information. This new alternative to charge-based systems, provides the possibility of achieving an ultra-low voltage operation and easier demonstration of digital systems coming from the bistable nature of spin [16].

In all-spin logic devices, input and output data are represented by the magnetization of two ferromagnets [17] which are communicating through a spin-coherent metallic channel. The physical view of these devices is shown in Fig 1(a). As shown in the Fig. 1(b-d) and discussed in [18], the applied voltage on the input ferromagnet, creates a flow of electrons which moves them from the supply voltage to the ground. This flow of electrons, becomes spin-polarized when passing through the input ferromagnet. Since the concentration of the spin-polarized electrons are different at the input side and the output side of the channel, the electrons diffuse to the output side. The accumulated spin-polarized electrons under the output ferromagnet, can switch the magnetization orientation of the magnet by applying a torque based on spin-transfer torque effect.

As shown in [19], these devices can be concatenable, exhibit nonlinear characteristics and support all Boolean operations. In all-spin logic operation, by using the direct spin signal, the nanomagnet can be switched and this signal can be transferred to the next stage. By storing the information in the spin magnetization of magnets, the input and output magnets can effectively be considered as digital capacitors linked by a spin-coherent channel. The sign and magnitude of control voltages applied on the magnets, determines the polarity of majority spin electrons and the device speed, respectively. Any change of magnetization in the bistable input magnet can exert a spin current through the channel and this current can determine the spin magnetization of the output magnet [17]. The channel between the two magnets can be either a metal or a semiconductor [25]. In our modeling and simulations we assume a copper interconnect.

The models utilized in this work are based on [17] where the different physical effects are captured. The accurate parameters of channel, magnet and interface that determine different performance characteristics, e.g., the spin injection, detection and transport efficiency are taken into account. The main important size effect parameters for the purpose of this work are the side wall specularity, the grain boundary reflectivity and the average grain size [18]. The average grain size is assumed to be equal to the width of thickness of the metals [18]. The complete list of parameters is in Appendix B.

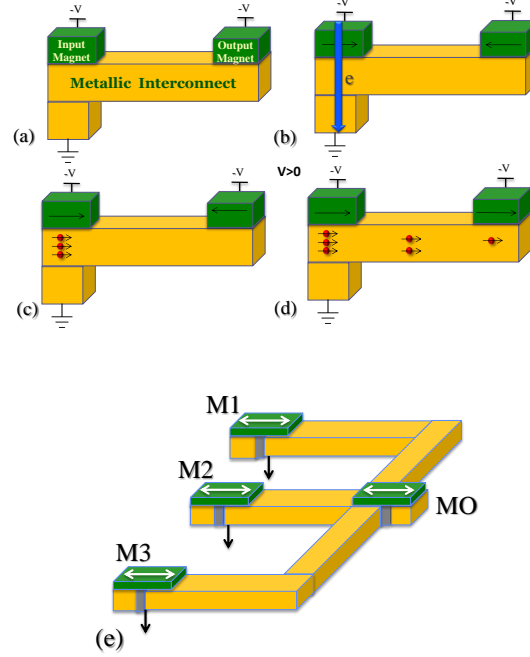


Figure 1: (a) Configuration of single ASL device (b) Applied voltage on the magnet, creates an electric field and enforces electron movements. (c) Spin-polarized electrons at the input side exhibit a higher density compared to the output side. (d) The diffusion of spin-polarized electrons towards the output magnet, changes the output magnetization direction. (e) An ASL Majority Gate with 3 inputs [17]. The 3 input magnets, M1, M2 and M3 are connected to the output magnet MO using 3 metallic interconnects.

2.1 Majority Gate Operation

As mentioned earlier, the ASL device supports a majority operation as shown in Fig 1(e). This feature is achieved because the net spin current to the output magnet can be determined by the sum of all input spin currents from all the input devices. In principle, this system can be designed for large number of inputs. As a trade off, by increasing the number of input devices in a majority gate, the uncorrelated thermal noise of these devices add up and impact the transient magnetization of output magnet, thus we need to make sure that in the design we have the proper fan-in. As it will be discussed later, if we only want to monitor the final steady value of the output magnetization, we can keep increasing the number of input devices as far as the output magnetization is predictable. Based on the device properties, this phenomenon sets a practical limit on the number of input devices to

a majority gate. On the other hand, if we care about the transient behavior of the output magnetization, fewer inputs should be connected to the output magnet to avoid the noise accumulation of input devices. In our simulations, for 3 and 5 input cases, the transient output magnetization is less impacted by the thermal noise, compared to higher fan-in numbers. We have to clarify that the steady state value of the majority gate depends on the sign of applied voltage on the magnets. In case of having a negative voltage applied on the magnets, the magnetization orientation value will finally be the exact majority of the input magnetizations. However, if the applied voltage is positive, the steady state value of the output magnet will be the complementary majority of the input magnetizations.

The interesting phenomenon in ASL majority gates rises from the dependency of the transient behavior of the output magnetization on the number of similar input magnetizations. This effect can be validated by the fact that the transferred spin torque increases when there are more magnets with magnetization in the same direction. Fig. 2, shows the different scenarios of transient output magnetization in majority gates with 3 and 5 inputs. As shown in Fig. 2(a), in a majority gate, with 5 inputs, the switching of output magnetization becomes faster when there are more inputs with similar magnetization directions. As the number of magnets with similar magnetization decreases, the switching happens slower and the effect of thermal noise is sensed more. In Fig. 2(b), the switching transition for two majority gates with 3 inputs and 5 inputs are compared. By considering the fact that the thermal noise accumulation in the gate with 3 inputs is less compared to the gate with 5 inputs, in the case of having equal net spin currents to both gates, the gate with 3 inputs, exhibits more deterministic transition.

2.2 Switching Delay Variation

The switching time of a ferromagnet is calculated in [30] using the small cone-angle approximation

$$\tau_{s,\omega} = \tau_0 \frac{\ln(\pi/\theta_0)}{\chi - 1}, \quad (1)$$

where τ_0 is a fitting parameter, θ_0 is the initial angle of the magnet, and χ is the ratio of the magnitude of injected spin current to the critical spin current required for the switching of the magnetization of the output ferromagnet. Based on this equation, if the value of injected spin current increases, the switching delay decreases. However, as shown in [17], the channel in this device can be approximated as an RC network; hence, the injected spin current and the value of supply voltage are directly correlated. Therefore, the switching delay is inversely proportional to the value of supply voltage. This result is shown in Fig. 3. The device parameters used for these simulations are shown in Appendix B.

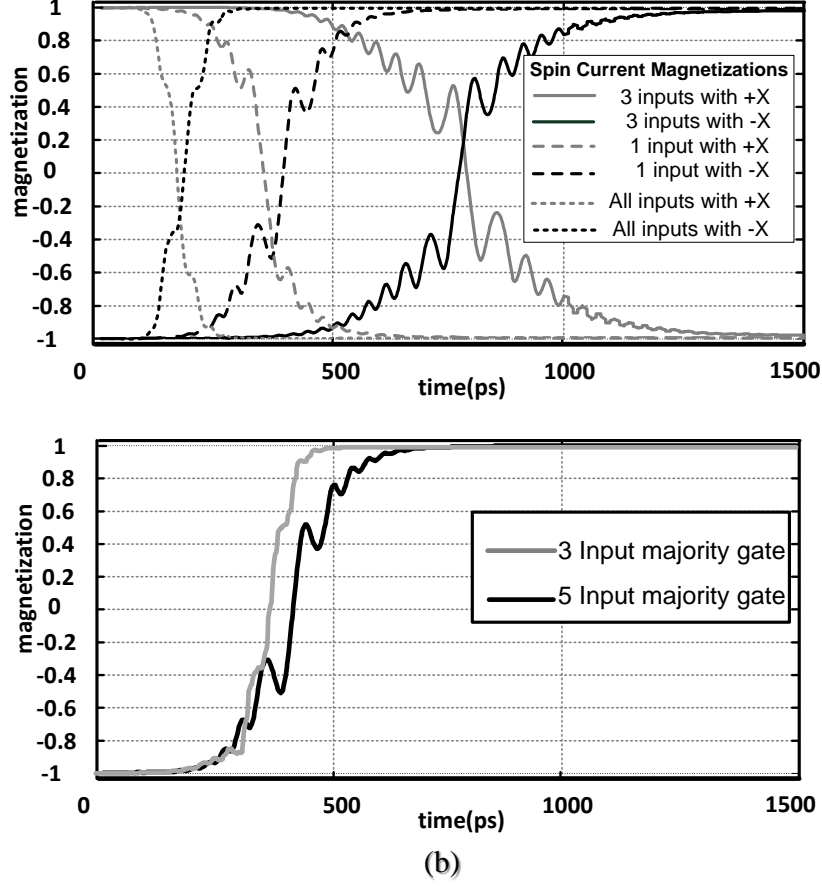


Figure 2: (a) Switching transient response for different scenarios of input magnetization in a majority gate with 5 inputs. (b) Switching transition comparison of majority gates with 3 and 5 inputs. In this comparison, the input magnetization of magnets to the 3 input gate are all similar. For the gate with 5 inputs, 4 inputs have similar magnetization and the net spin current is equal to the other gate. The applied voltage on the magnets in these simulations is -5 mV.

2.3 Impact of Thermal Noise

The thermal motion of electrons inside the ferromagnets, is the main cause of thermal noise. The amount of noise is correlated to the temperature of magnets and can directly

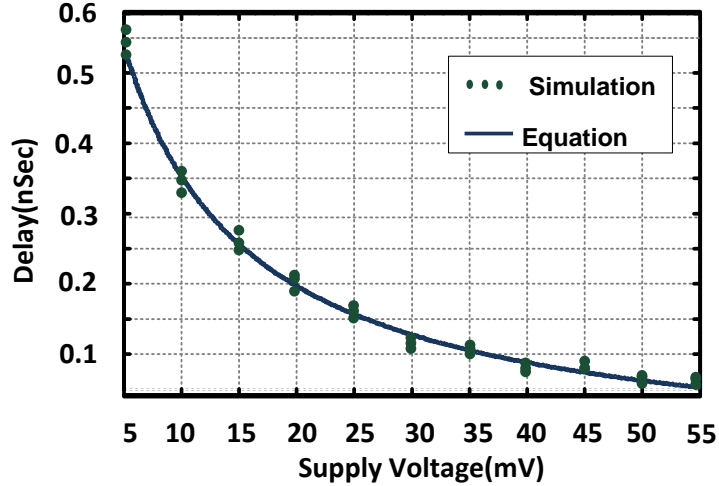


Figure 3: Switching delay variation versus the supply voltage. Each point is simulated 3 times to verify the results.[22]

affect the steady-state precession angel θ_0 . Based on the derivations in [29]

$$\langle \theta_0^2 \rangle = \frac{K_b T}{E_b}. \quad (2)$$

In this equation, E_b is the barrier energy and T represents the temperature. This thermal effect acts as the main source of noise which can impact the output magnetization. In our simulations, θ_0 can differ 5% to 10% from the analytic solution based on different parameters.

3 Pattern Recognition Scheme

Similar to any recognition system, in this work we consider two major phases for the operation. The first phase is the “learning” phase where the desired pattern is stored in memory. In “detection” phase, the circuit identifies the similarity of an input data and the stored pattern with respect to the decision making criteria. In the learning phase, the circuit can receive a single image or a training set. The training set includes multiple training images from different users.

In this section, we propose a new technique using all-spin logic devices and establish a fully spin-based operation. By illustrating several examples, we verify the performance for various image sizes.

3.1 Mainly Similar Images

We first provide the mathematical definition of mainly similarity and then show how this can help the training of the circuit. In our simulations, all the images are binary-valued matrices with 0 and 1 representing white and black pixels, respectively. In our circuit, we assume that binary “0” logic corresponds to magnetization orientation in $-X$ direction and binary “1” logic corresponds to magnetization orientation in $+X$ direction.

For a given pair of binary vectors x and y with equal length L , the *Hamming distance* [21] is defined as

$$d(x, y) = \sum_{i=1}^L 1 - \delta_{x_i y_i},$$

where x_i and y_i denote the i^{th} components of x and y respectively and δ is the kronecker delta. Subsequently, we can exploit this quantity as a measure of similarity between two images.

Definition 1 *Two binary images B and $B' \in \{0, 1\}^{m \times n}$ are called mainly similar if the majority of pixels across every two rows are identical. More specifically,*

$$\forall k \in \{1, \dots, m\} : \quad d(B_{k,:}, B'_{k,:}) < \lfloor \frac{n}{2} \rfloor, \quad (3)$$

where $B_{k,:}$ denotes the k^{th} row of B and $\lfloor a \rfloor$ represents the floor operation on a (i.e., the largest integer not greater than a).

By this comparison, we ensure that the two images have almost similar pixels along the corresponding rows. For the purpose of this paper, we considered the comparison along the rows, although a column-wise comparison could be established with no loss of generality. As illustrated in Fig. 4, being mainly similar along the rows, does not imply being similar along the columns.

3.2 Majority Training and Decision Making

In the learning phase, we train the circuit by providing a number of mainly similar images. In reality, these images could be different representations of a target image (say a character or a certain binary pattern). We build up a representative of the given similar images by constructing a so-called *mean image*.

Definition 2 *For a set of P binary images $B_1, B_2, \dots, B_P \in \{0, 1\}^{m \times n}$, the corresponding mean image denoted as \bar{B} is a binary image with entries*

$$\bar{B}(i, j) = \text{nint}\left(\frac{1}{P} \sum_{k=1}^P B_k(i, j)\right). \quad (4)$$

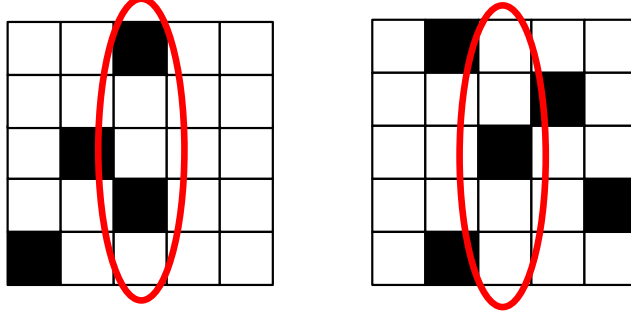


Figure 4: The two images are mainly similar (along the rows), however, the Hamming distance between the third columns is 3 which does not imply a similarity along the columns

In this equation, $nint$ denotes the nearest integer function. In our circuit, the mean image represents the desired pattern by the users and is utilized as a reference. Since this matrix is constructed using all-spin majority gates, the number of training images, P , is considered to be odd and upper bounded by the maximum number of inputs to a majority gate as discussed in subsection 2.1.

After the training data is stored and the mean image is constructed, we make a row-wise comparison between the input and the mean image. As we will see in the next section, depending on the initial value of output magnetization, the non-Boolean row decision maker can return the total count of matches or mismatches between the compared rows of input image and the mean image.

4 Proposed Structure and Design Considerations

Based on the pattern recognition scheme shown in Section III, we study two different implementations of the circuit. By comparing the performances of the two different versions of the *single pixel comparator* unit, we choose the one with more capabilities, at the expense of slightly more power consumption and occupied area. In the single pixel comparator, the circuit receives the training pixels from P different users and the mean image is constructed, subsequently. The value of the mean pixel is then compared with the corresponding value in the input image and the steady state magnetization of *Pixel* magnet stores this information. The two versions of this unit both operate based on the idea of training the circuit with a set of mainly similar images and comparison of the single pixels from the input image with their correspondence in the mean image. With respect to

the required operations, the single pixel comparator, needs a memory to store the training data, a logic comparator and a circuit to construct the mean pixel. As previously mentioned, the mean pixel can be constructed by an all-spin majority gate; however, for the memory and the comparator, we will propose a new circuit in the following subsection.

4.1 Memory+Logic Comparator

1-bit full adder structures with a total number of 5 nanomagnets have been designed in [22] and [24]. By proper setting of the circuit in [22], we use it as an area and power efficient comparator (XNOR) block as shown in Fig 5. The two inputs to this block (A

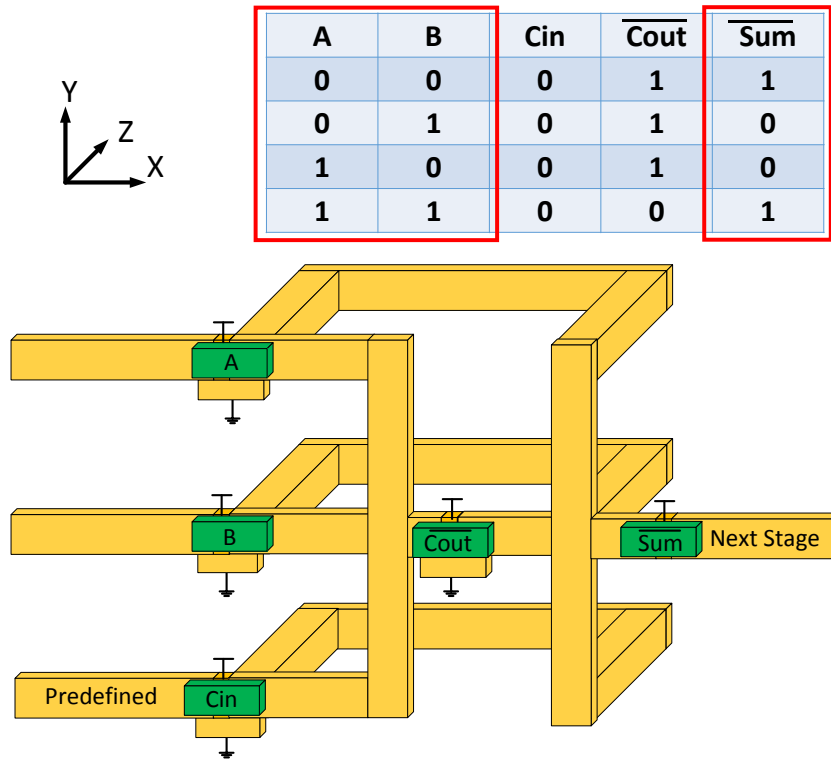


Figure 5: 1-bit full adder used as XNOR[22]. In the 2D implementation of this work, X and Y wires are in-plane metal wires and connections along the Z axis are vias.

and B) are coming from distinct sources. One of the inputs comes from the input image synchronized with the control voltage and the other input is given to the circuit during the learning phase. Compared to a CMOS counterpart, this structure exhibits very important

advantages. First, it requires 5 magnets whereas the CMOS version requires at least 8 transistors for XNOR implementation. Second, this circuit has the capability of storing the training information without extra static power consumption, whereas in CMOS, excess power is consumed to store this data [10]. Taking advantage of the non-volatile operation in ASL devices, the input magnets of this circuit can store the binary data and later the stored information is used to determine the magnetization direction of the next stages. Fig. 6 shows the simulated output waveform (*sum* magnet) of the XNOR block for different scenarios of input magnetization. As it is important to consider the breakdown current effects [23], we choose the 5mV supply voltage in our simulations. This is to ensure that the current density is safely below the breakdown value. It is noteworthy that for channels with higher breakdown current densities, higher voltages can be applied and the operation speed increases. In this simulation, a control voltage of 5mV is applied on the magnets at $t = 0$. The total power consumption of the XNOR gate is $11\mu W$ and the estimated area is less than $0.3\mu m^2$. As we apply a control voltage on the XNOR gate, the output magnetization remains in $-X$ orientation (initial condition of magnetization in this simulation) if the pixel values are different. In case of having similar inputs to this gate, the output magnetization switches to $+X$ direction as shown in Fig. 6. We have to clarify that the initial condition of output magnet does not change the final magnetization orientation.

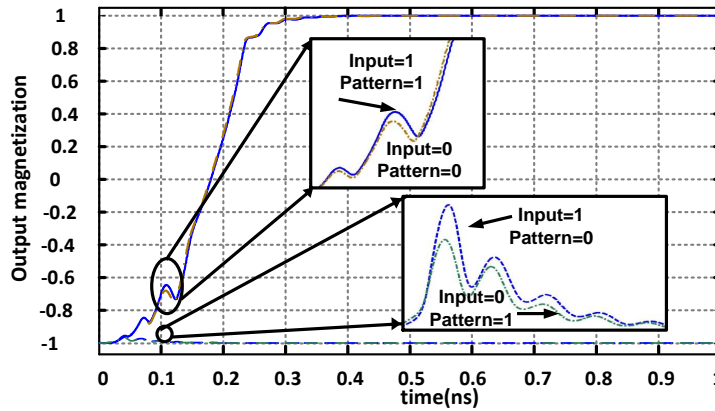


Figure 6: Simulated output waveforms of XNOR gate

4.2 Construction of the mean pixel

As a reliable and simple way to extract the information from the training set, we construct the mean image as discussed in Section III. The ASL majority gate with the schematic shown in Fig. 1(b), provides a low power and efficient implementation of the mean image.

The inputs to this majority gate, come from P different users and the images that system receives during the learning phase are constrained to be mainly similar along the rows. By applying the control voltage on the magnets of this gate, the output magnetization either switches to other value or remains in the same magnetization orientation. If the applied control voltage is negative, the output final magnetization orientation is the majority of the input magnetizations. In the case of positive control voltage, the output magnetization settles to the complementary majority value of the input magnetizations. For this system, since we apply unified positive voltages, the majority gates settle to the complementary majority value. In order to extract more information from the majority gates operation in this circuit, we assume a unified value of initial magnetization orientation on the output magnets of each stage of majority gates. This enables us to recognize the total count of matches or mismatches between the input magnetizations to each majority gate, as we will discuss later. The total power consumption of each majority gate in this circuit is $3.75\mu W$ and the estimated area is less than $0.2\mu m^2$.

4.3 Single Pixel Comparator

By having the required blocks, we propose the two different versions of the single pixel comparator.

4.3.1 Standard implementation

The schematic of this implementation and the table with the detailed operation are shown in Fig. 7. This circuit operates in the same order discussed in Section III. The first stage of the circuit is a majority gate with inputs coming from the P users in the learning phase. The output of this majority gate settles to the corresponding mean pixel value. The output of this gate is connected to a comparator circuit which has the other input coming from the input image. The connection is through a short metallic interconnect to minimize the delay. When the learning phase is over and the detection phase starts, by applying the control voltage across the magnets of the comparator circuit, the “Pixel” magnet settles to the comparison value of the mean pixel and the input pixel. It is noteworthy that the input pixel can be applied on magnet Q_{ij} after the \bar{P}_{ij} magnetization settles to the mean pixel; hence, no extra memory circuit is required to store the value of \bar{P}_{ij} .

4.3.2 Comparator-First implementation

In this version, there will be the same number of comparator circuits as the total number of training images at the input side. The comparators have the input image pixel, Q_{ij} in common and differ in their other input which comes from their corresponding training image. The output magnets of the comparators are connected to the Pixel magnet through

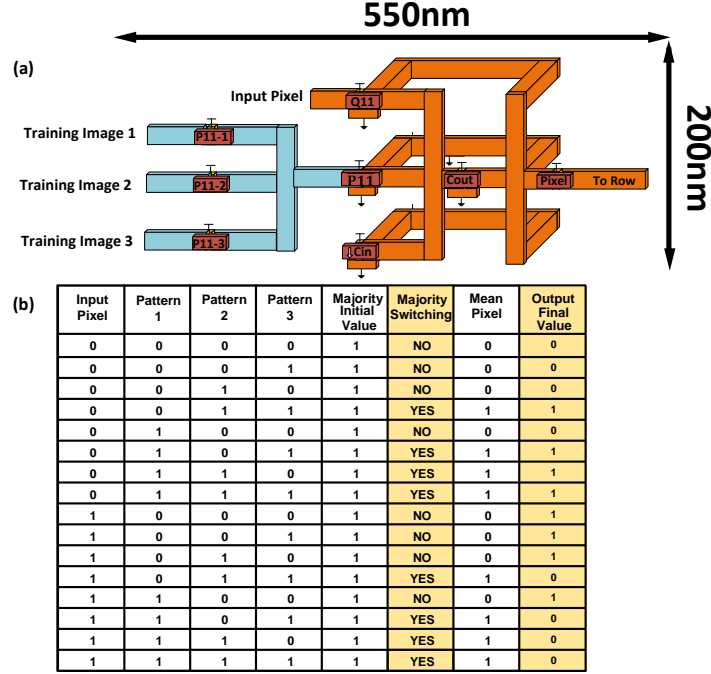


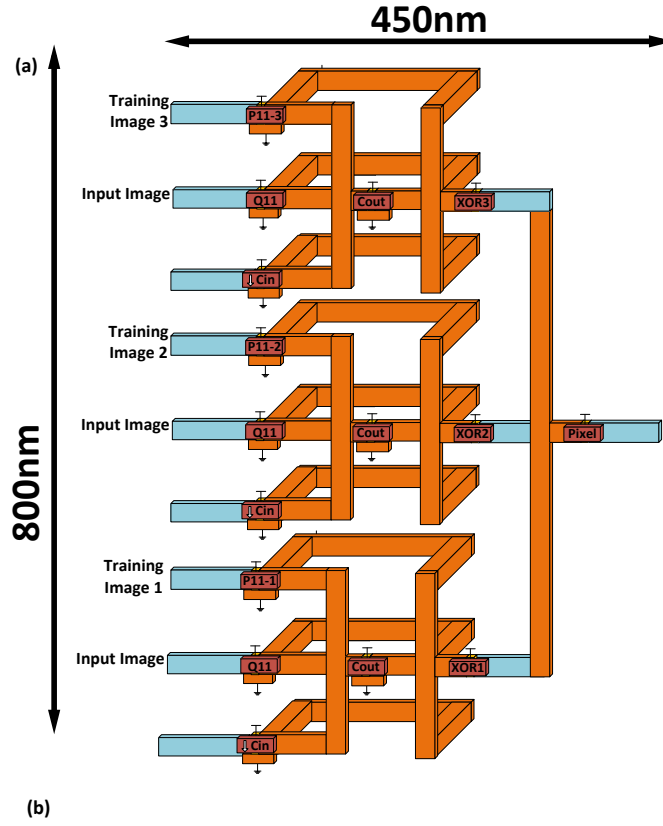
Figure 7: (a) Standard single pixel detector schematic. (b) The truth table with the detailed operation of the circuit.

metallic interconnects in a majority gate configuration. During the learning phase, the pattern pixels are stored in the corresponding input magnets. By applying the control voltage on the magnets of the circuit, the detection phase starts and the “Pixel” magnetization settles to the comparison value of the mean pixel and the input pixel. The schematic of the circuit and the detailed operation table are shown in Fig. 8.

As it can be verified by comparing the last columns of Fig. 7(b) and Fig. 8(b), the “Pixel” steady state value is identical in the two versions. To verify the identical output result from the two different versions of the implementation in a more general case, we have to prove that the majority operation and the comparison (XOR/XNOR) operation are interchangeable, i.e.,

Proposition 1 *Given x, y_1, y_2, \dots, y_P as binary variables and P as an odd integer number,*

$$x \oplus \text{uint}\left(\frac{1}{P} \sum_{k=1}^P y_k\right) = \text{uint}\left(\frac{1}{P} \sum_{k=1}^P (x \oplus y_k)\right), \quad (5)$$



(b)

Input Pixel	Pattern 1	Pattern 2	Pattern 3	Output Initial Value	Output Switching	Output Final Value
0	0	0	0	1	YES	0
0	0	0	1	1	YES	0
0	0	1	0	1	YES	0
0	0	1	1	1	NO	1
0	1	0	0	1	YES	0
0	1	0	1	1	NO	1
0	1	1	0	1	NO	1
0	1	1	1	1	NO	1
1	0	0	0	1	NO	1
1	0	0	1	1	NO	1
1	0	1	0	1	NO	1
1	0	1	1	1	YES	0
1	1	0	0	1	NO	1
1	1	0	1	1	YES	0
1	1	1	0	1	YES	0
1	1	1	1	1	YES	0

Figure 8: (a) Comparator-first pixel detector schematic. (b) The truth table with the detailed operation of the circuit.

where \oplus denotes the XOR operation. The mathematical proof of this proposition is shown in Appendix A.

Although the standard implementation has slightly lower power consumption (Less number of devices) and a smaller area, we select the comparator-first design as the unit cell of this circuit. This is due to the fact that the output magnetization transient of this circuit provides more information on the similarity of the training pixels and the input pixel. Based on Fig 7(b) and Fig 8(b), the final value of output magnetizations, in the two cases are identical. However, the Comparator-first output magnetizations is coming from a majority gate and switches when the majority of pattern pixels have the same value of the input pixel. If the majority gate at the output of Comparator-first circuit has a low fan-in (e.g., ≤ 5), the switching transient behavior will be less sensitive to the accumulated thermal noise and the information on the number of training pixels with identical values will be provided. On the other hand, in the standard implementation, the output magnetization is from the XNOR circuit and conveys no information on the number of similar pattern pixels. Based on Fig 7(b), the output magnetization transient will not add information on the number of training pixels with identical values. This is particularly important when the user in the detection phase tracks the total count of pattern pixels with the similar value.

4.4 Non-Boolean Row Decision-Maker

The last stage of the proposed circuit uses the interesting feature of the ASL majority gate as a means to quickly decide about the mainly similarity of the input image and the mean image, along the rows. The inputs to this majority gate is from the “Pixel” magnets of the pixels along the same row of the image. The connection is through short interconnects to minimize the delay. As mentioned before, the spin torque transferred from the input magnet to the output magnet in the ASL majority gate, is determined by the magnetization of the input devices. As the number of devices with similar magnetization orientations increases, the transferred spin torque increases; hence, the output magnetization switching becomes faster according to (3). By proper selection of the control voltage timing and also the dimensions of the nanomagnets and metallic interconnects in this gate, a reliable decision-making based on the transient behavior of output magnetization is achieved. This final majority gate is sensitive to the uncorrelated thermal noise of input magnets; hence, an intentional low fan-in number (≤ 5) has to be selected. In our simulations, 3 magnets from the previous pixel stages are connected to this gate and as it will be shown in simulation results, a reliable decision-making is achieved.

The complete circuit for the full image comparison consists of two stages. The unit pixel comparator and the row majority gate. The structure consisting of the comparator-

first circuits and the Row majority gate for a 3×3 pixel image comparison is called the “Smart Detector Cell”. This naming convention, helps the discussion of operation in the next section. We call these detector cells smart because they can perform multiple tasks of “storage”, “Boolean Computation” and “non-Boolean decision-making” in a time-efficient manner. The schematic of this circuit is shown in Fig. 9. The total power consumption of this circuit is $115 \mu W$ and the occupied area is less than $0.5 \mu m^2$.

We have to mention that in our simulations, we have taken into account the effect of magnetization switching. The ASL device acts like a resistive network and the power consumption will not change with time. In these devices, the current passes through only one magnet and therefore does not change with time and the magnet switching of the input side, will influence the switching delay of the last magnet. In this paper, we consider the worst case delay which takes into account the switching delay from the input magnet of the first device to the output magnet of the last device as well as the transport delay within the metallic interconnects. For DC power consumption estimations, we have performed DC and transient simulations and the results are consistent. It is noteworthy that the low operational voltage of the circuit will lead to a low power consumption.

In a real implementation of this work, read/write circuits are added to fully realize the circuit. However, this paper focuses on the processing circuit without concerns regarding the feeding and extraction of the input and output data. In order to feed the input data, spin polarized currents are used to initialize the magnetization of input magnets based on the training images, similar to [12]. On the other hand, the number of write units is equal to the number of pixels, while there is one output, which translates to small overhead. The decision data is in form of time delay and can be stored on a capacitor, where the delay impacts the amount of the stored charge. The other possibility to extract the output data will be using MTJ devices, as mentioned in [33].

5 Simulation Results

In this section, we provide two different examples to show the reliable performance of smart detector cells.

5.1 Non-Boolean Hamming Distance Identifier of 3×3 Pixel Pattern and Input Image

In this example, we only have one training image and one input image. To compare the similarity of these two images, we need 9 XNOR gates to identify the similarity of corresponding pixels in the two images and 3 majority gates with Fan-in of 3 to decide on the similarity of the corresponding rows. It is also obvious that the mean image in this case will be the same pattern image. The smart detector cell in Fig. 9, has 3 comparator-first circuits and a *Row* majority gate. The mainly similarity of the rows can be determined by

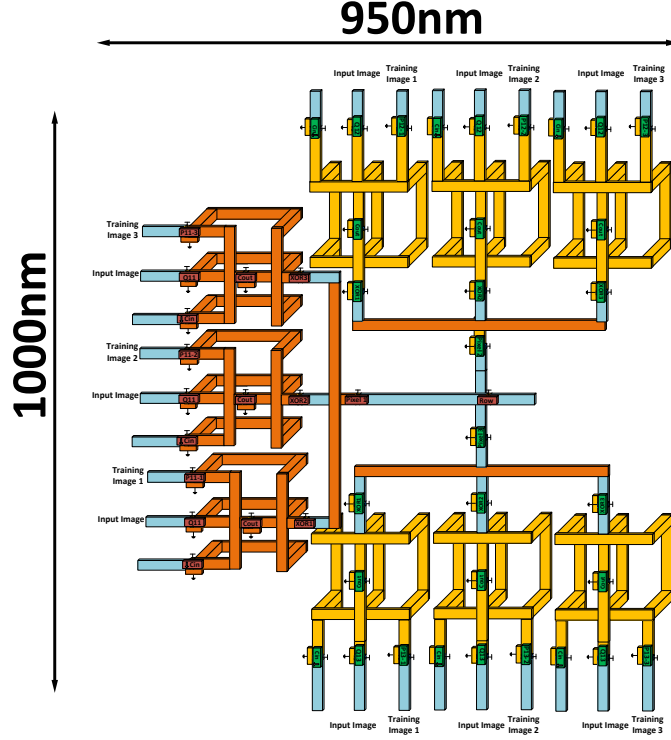


Figure 9: Structure of the unit smart detector cell

the Pixel majority gates. The last majority gate in this case, settles to $+X$ magnetization if at least 2 rows are mainly similar. However, in this simulation, the output magnetization of this gate is not important. In this simulation, the initial magnetizations of the comparators and the majority gate outputs are set to 1. Fig. 10 shows the two images as well as the transient magnetization for various magnets. The Pixel waveforms overlap in some cases and that is why we are showing only 3 pixels in this figure. As expected, the comparator outputs switch for P_{21} and P_{22} pixels since the values in the input image and the pattern image are different. For the rest of pixels, the comparator output is $+X$ magnetization and will not switch. Subsequently, row 1 and row 3 both exhibit perfect similarity and the output of the corresponding majority gates switch within the shortest time as it can be compared with Fig. 2(b). On the other hand, row 2 exhibits a mismatch and therefore can not switch to $-X$ magnetization orientation. The control voltage of 5 mV is applied on all the magnets at $t = 0$ and the circuit compares the two images in less than 0.6 ns. Compared to CMOS circuits, this exhibits a much lower operational voltage and decision time.

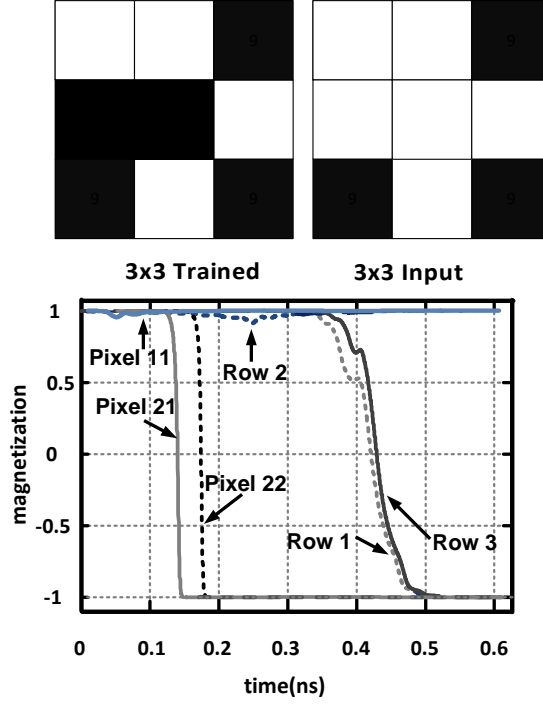


Figure 10: Using a single smart detector cell, we can compare these 3×3 pixel images. The waveforms of the comparators and majority gates (bottom).

5.2 Non-Boolean Similarity Comparison of a 9×9 Pixel Image and a Set of 3 Pattern Images

In order to incorporate the smart detector cells for larger images, we need an accurate design of the cells. Here, we develop a circuit for training with 9×9 pixel images and perform a non-Boolean comparison between the constructed mean image and the 9×9 pixel input image. In this simulation, 3 different users write the word “Spin” by their own choice of pixels. The 3 pattern images are shown in Fig. 11.

In the detection phase, a new user of the circuit, chooses an arbitrary image of interest as the input. As an example in this simulation, the user chooses the word “swim” as shown in Fig. 12 (left). The circuit should compare this image and the mean image constructed from the training set.

The mean image of the training set is also shown in Fig. 12 (right). One particular

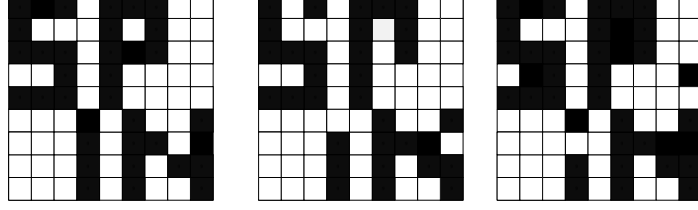


Figure 11: Training set for the 9×9 pixel images

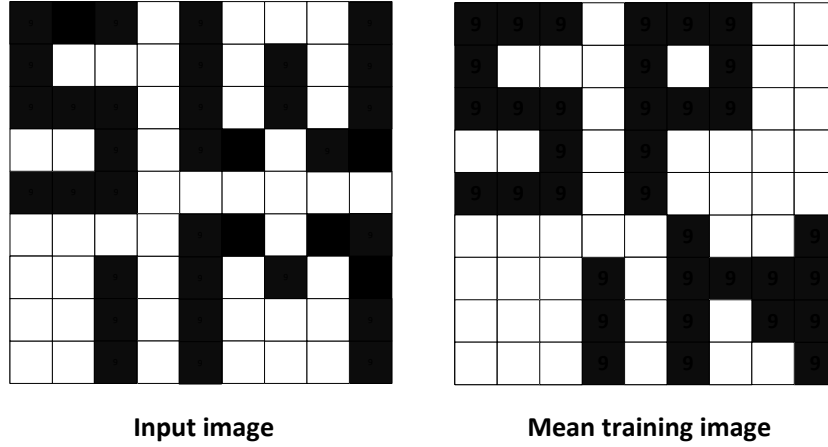


Figure 12: The input image (left) and the mean image (right).

advantage of constructing the mean image can be discussed here. As it can be seen in the mean image, those pixels which are mistakenly valued by a single user (e.g., P_{26} and P_{49}) in the learning phase, are automatically corrected when the mean image is constructed. This is specifically useful, when the users during the learning phase, train the system with multiple versions of an image to make sure that the mean image represents their desired pattern. The mistaken values could be due to any source of error or distortion. In an ideal case where the thermal noise effect can be ignored, by simply changing the fan-in of different stages in the smart detector cell, the circuit can compare these two large images. However, in our simulations, as we model the thermal noise accurately, the fan-in considerations mentioned before, are particularly important. Based on these considerations, we break these 9×9 images into smaller 3×3 subimages, where a single smart detector cell

unit can be used for the comparison. The 9 smart detector cells can operate in parallel and the circuit configuration can be determined by the user. By this breakdown, we can also achieve more information on the pixels as we can check the mainly similarity for smaller blocks of the original image. The breakdowns of the mean image (squares on the right) and the input image (squares on the left) are shown in 3×3 partitions in Fig. 13.

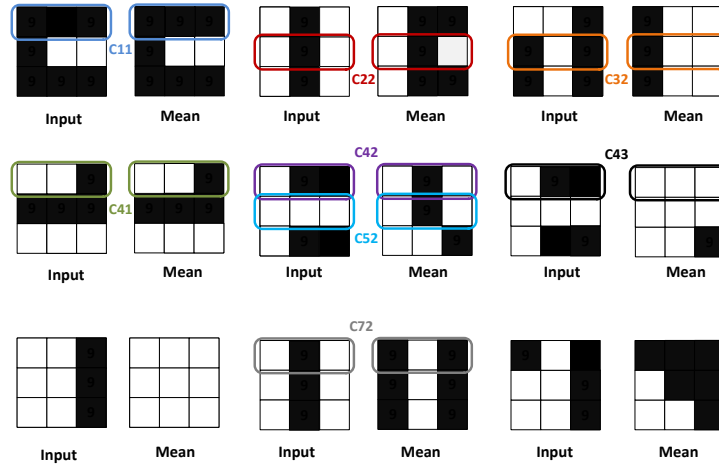


Figure 13: Due to fan-in considerations, the circuit is consisted of 9 smart detector cells. The corresponding breakdowns of the mean image and the input image are shown here .

In order to distinguish the different rows of smaller blocks, we use the notation of C_{ij} clusters, which represents the elements of the i^{th} row from column $3j - 2$ to column $3j$. The magnetization waveforms shown in Fig. 14 and Fig. 15 separately show the output magnetizations of smart detector cells for various clusters. The unified initial condition of the output magnet in this simulation is $-X$ magnetization orientation. In Fig. 14, the switching delay of output magnetizations for the clusters with perfect match (C_{11} , C_{22} and C_{41}) and those with 1 mismatch (C_{52} , C_{42} and C_{32}) can be easily distinguished. This phenomenon was previously described as the unique feature of ASL majority gates and helps the users to identify the number of mismatches along different rows. At the same time, the output magnetization of the clusters with the same level of similarity, are very close in time domain which makes this non-Boolean decision-making a reliable metric. On the other hand, in Fig. 15, the output magnetization can not switch for clusters with mismatches (C_{43} and C_{72}) and as it can be seen, the level of precession for different mismatch levels is not the same. This is due to the different amount of spin torques provided in these two cases. If the user has a very high resolution study on the output magnetization, this can help to identify the number of mismatches; however, the switching transient is a more

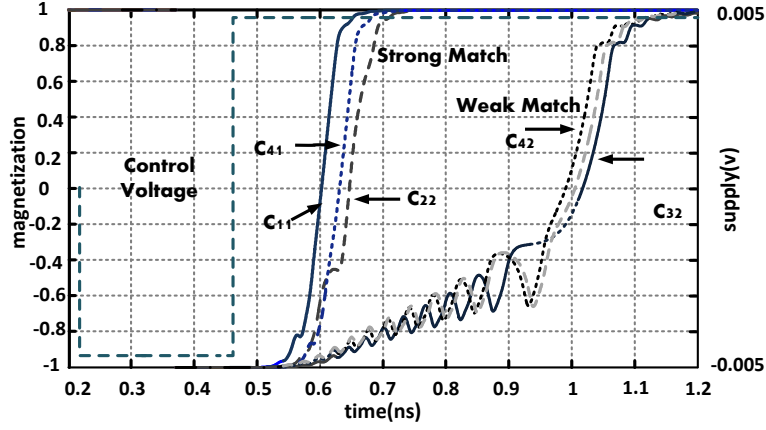


Figure 14: The switching delay of output magnetization in last stage represents the similarity of input data and pattern data.

reliable metric and the same information can be extracted by repeating the simulation with the output magnet initial condition set to $+X$ magnetization.

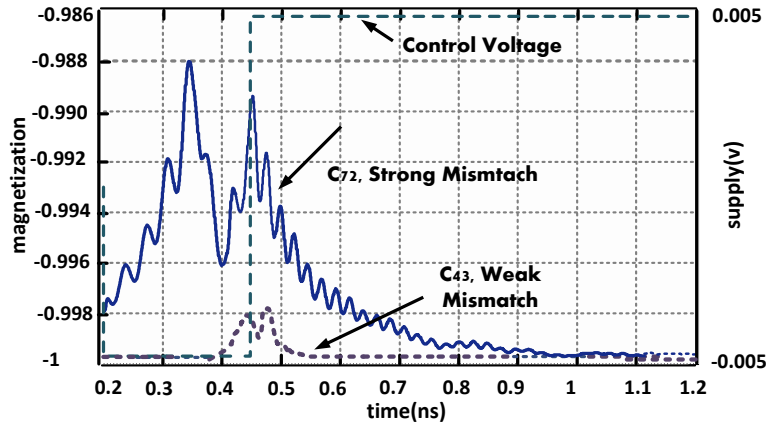


Figure 15: Since these clusters represent mismatch, they can not switch and the initial magnetization does not change. Note that the y-axis is showing from -1.002 to -0.986 in contrast with Fig. 14 in which the y axis is from -1 to 1.

Table 1: Performance Comparison with existing CMOS systems

Reference	[7]	[12]	This Work
Decision time	30ns	N.A.	1 ns
Image Size	32×32	86 neurons	9×9
DC Power	N.A.	$2.2mW$	990 uW
Area	N.A.	$0.018mm^2$	$< 1\mu m^2$
Technology	CMOS	Spin-CMOS	All-spin

As it can be seen in all the simulation results, this circuit can make a decision in almost 1000 ps for a 9×9 pixel image, whereas in CMOS, this decision time, can not be less than few nanoseconds. For a detailed comparison between the two technologies, in table 1, the performance of this circuit and few existing CMOS circuits are compared.

6 Conclusion

We have presented a novel non-Boolean image recognition circuit based on all-spin logic devices. The introduced circuit can perform all the phases of a non-Boolean pattern recognition for binary images. Taking advantage of the non-volatility of ASL devices, the learning phase operation is performed incorporating no additional memory devices. By introducing the mainly similarity scheme, two different implementations of the circuit are proposed. As verified by simulation results, this circuit can recognize various sizes of binary image patterns faster than existing CMOS counterparts and consumes less power with an operational voltage of 5mV. Since the comparisons in this circuit are based on ASL majority gates, the computational complexity of the operation is also less compared to existing circuits. The proposed circuit has applications in fast and low power image recognition for security, medical imaging, and sensing.

Appendix A: Proof of proposition 1

In this appendix, we mathematically verify (7). Since all the variables are binary-valued,

$$0 \leq \frac{1}{P} \sum_{k=1}^P y_k \leq 1.$$

The *nint* operation results in 0, when

$$\sum_{k=1}^P y_k < \frac{P}{2}, \quad (6)$$

otherwise it results in 1. Therefore, there are 4 different possibilities for the variables, as shown in table II. In order to simplify the notations, we also define,

$$z_k = x \oplus y_k \quad \forall k \in \{1, \dots, P\}. \quad (7)$$

Table 2: Possibilities of x, y_1, \dots, y_P

x	$\sum_{k=1}^P y_k$	$nint(\frac{1}{P} \sum_{k=1}^P y_k)$	$x \oplus nint(\frac{1}{P} \sum_{k=1}^P y_k)$
0	$< \frac{P}{2}$	0	0
0	$> \frac{P}{2}$	1	1
1	$< \frac{P}{2}$	0	1
1	$> \frac{P}{2}$	1	0

Here, we verify (7) for the first row of table II. Using the same method for the other 3 rows, the proposition can be completely proved.

If $\sum_{k=1}^P y_k < \frac{P}{2}$, fewer than $\frac{P}{2}$ of y_k 's are 1. Given $x = 0$, this means that fewer than $\frac{P}{2}$ of z_k 's are 1 and the rest are zero, i.e.,

$$\sum_{k=1}^P z_k < \frac{P}{2}.$$

Similar to (8), by applying the nearest integer function,

$$nint(\frac{1}{P} \sum_{k=1}^P z_k) = 0.$$

Appendix B: Simulation Parameters

<i>Size Effect Parameters</i> [18]		
Side Wall specularity	P	0
Grain Boundary Reflectivity	R	0.2
<i>Interface Parameters</i> (Co/Cu)[28]		
Majority Spin conductance	G_{\uparrow}	$0.375 \text{ } 1/\Omega$
Minority Spin conductance	G_{\downarrow}	$0.125 \text{ } 1/\Omega$
Real Spin-Mixing Conductance	$\text{Re}G_{\uparrow\downarrow}$	$3.43751/\Omega$
Imaginary Spin-Mixing Conductance	$\text{Im}G_{\uparrow\downarrow}$	$9.37 \times 10^{-3} \text{ } 1/\Omega$
<i>Ferromagnet</i> (Co)[31]		
Ferromagnet Length	L_x	75.00 nm
Ferromagnet Width	L_y	25.00 nm
Ferromagnet Height	L_z	3.00 nm
Gilbert Damping Coefficient	α	0.0021
Gyromagnetic Ratio	γ	$1.76 \times 10^{11} \text{ } 1/\text{sT}$
Saturation Magnetization	M_s	$1.45 \times 10^6 \text{ A/m}$
Number of spins in magnet	N_s	$1.34 \times 10^6 \text{ } 1/\text{V}$
Energy Density	K_u	$0.5 \times 10^5 \text{ J/m}^2$
<i>Channel</i> (Cu)[32]		
Channel Length	L_{int}	212.5 nm
Channel Width	W_{int}	50 nm
Thickness/Width aspect ratio	AR	2.0
Channel Thickness	H_{int}	100.0 nm
Cross section Area	A	5000 nm^2
Finite difference spacing	Δx	10.0 nm
Conductivity	σ	$41.549 \text{ } 1/\mu\Omega\text{m}^2$
Diffusion coefficient	D	$0.014 \text{ m}^2/\text{s}$
Permeability	μ	$0.003 \text{ m}^2/\text{Vs}$
Spin relaxation time	τ_s	10.939 ps

Acknowledgment

The authors would like to thank Mr. Vahnood Pourahmad, Dr. Abolhassan Vaezi and Dr. Mohammad Emadi from Cornell University and Dr. Alireza Aghasi from Georgia Institute of technology for helpful discussions. We also, like to thank Semiconductor Research Corporation (SRC) and Institute for Nanoelectronics Discovery and Exploration (INDEX) for the financial support of this project.

References

- [1] S. Wold. “Pattern recognition by means of disjoint principal components models.” *Pattern recognition* 8.3 : 127-139, 1976.
- [2] J. Flusser, and T. Suk. “Pattern recognition by affine moment invariants.” *Pattern recognition* 26.1 : 167-174, 1993.
- [3] DH. Ballard “Generalizing the Hough transform to detect arbitrary shapes.” *Pattern recognition* 13.2 : 111-122, 1981.
- [4] J. S. Seo, B. Brezzo, Y. Liu, B. D. Parker, S. K. Esser, R. K. Montoye ... and D. J. Friedman “A 45nm CMOS neuromorphic chip with a scalable architecture for learning in networks of spiking neurons”. In *Custom Integrated Circuits Conference (CICC)*, IEEE (pp. 1-4). IEEE, 2011.
- [5] M. Koyanagi, Y. Nakagawa, K. W. Lee, T. Nakamura, Y. Yamada, K. Inamura, ... and H. Kurino “Neuromorphic vision chip fabricated using three-dimensional integration technology” In *Solid-State Circuits Conference, 2001*. Digest of Technical Papers. IEEE (pp. 270-271), 2001.
- [6] R. W. Hlzel and K. Krischer. “Pattern recognition with simple oscillating circuits.” *New Journal of Physics* 13.7 : 073031, 2011.
- [7] S. P. Levitan, Y. Fang, D. H. Dash, T. Shibata, D. E. Nikonov, and G. I. Bourianoff “Non-Boolean associative architectures based on nano-oscillators”. In *Cellular Nanoscale Networks and Their Applications (CNNA)*, 13th International Workshop on (pp. 1-6). IEEE, 2012
- [8] M. Ishikawa, K. Ogawa, T. Komuro and I. Ishii “A CMOS vision chip with SIMD processing element array for 1 ms image processing”. In *Solid-State Circuits Conference, 1999*. Digest of Technical Papers, (pp. 206-207). IEEE, 1999.
- [9] J. Liu and M. Brooke, “Fully parallel on-chip learning hardware neural network for real-time control” in *Proc. IEEE Int. Symp. Circuits Syst.*, vol. 5, pp. 371374, 1999
- [10] Matsunaga, Shoun, Jun Hayakawa, Shoji Ikeda, Katsuya Miura, Haruhiro Hasegawa, Tetsuo Endoh, Hideo Ohno, and Takahiro Hanyu. “Fabrication of a nonvolatile full adder based on logic-in-memory architecture using magnetic tunnel junctions.” *Applied Physics Express* 1, no. 9 (2008): 091301.
- [11] Dmitri E. Nikonov, and Ian Young. ”Overview of beyond-CMOS devices and a uniform methodology for their benchmarking.” *Proceedings of the IEEE* 101, no. 12 (2013): 2498-2533.

- [12] M. Sharad, C. Augustine, G. Panagopoulos, and K. Roy “Spin-based neuron model with domain-wall magnets as synapse” *Nanotechnology, IEEE Transactions on*, 11(4), 843-853.
- [13] V. Q. Diep, B. Sutton, B. Behin-Aein, and S. Datta (2014) “Spin switches for compact implementation of neuron and synapse. *Applied Physics Letters*” 104(22), 222405.
- [14] S. Datta, S. Salahuddin and B. Behin-Aein. “Non-volatile spin switch for Boolean and non-Boolean logic.” *Applied Physics Letters* 101.25 : 252411, 2012
- [15] L. Su, W. Zhao, Y. Zhang, D. Querlioz, Y. Zhang, J. Klein, P. Dollfus, and A. Bournel. “Proposal for a graphene-based all-spin logic gate.” *Applied Physics Letters* 106, no. 7 (2015): 072407.
- [16] Augustine, Charles, et al. “Ultra-low power nanomagnet-based computing: a system-level perspective.” *Nanotechnology, IEEE Transactions on* 10.4 : 778-788, 2011.
- [17] P. Bonhomme, S. Manipatruni, R.M. Iraei, S. Rakheja, Sou-Chi Chang, D.E. Nikonov, I.A. Young, A. Naeemi “Circuit Simulation of Magnetization Dynamics and Spin Transport,” *Electron Devices, IEEE Transactions on* , vol.61, no.5, pp.1553,1560, May 2014
- [18] R. M. Iraei, P. Bonhomme, N. Kani, S. Manipatruni, D. E. Nikonov, I. A. Young and A. Naeemi, “Impact of dimensional scaling and size effects on beyond CMOS All-Spin Logic interconnects”. In *Interconnect Technology Conference/Advanced Metallization Conference (IITC/AMC)*, 2014 IEEE International (pp. 353-356). IEEE, 2014.
- [19] B. Behin-Ain, D. Datta, S. Salahuddin, and S. Datta, “Proposal for an all-spin logic device with built-in memory, *Nature Nanotechnol.*, vol. 5, no. 4, pp. 266270, 2010
- [20] C. Augustine, G. Panagopoulos, B. Behin-Ain, S. Srinivasan, A. Sarkar, and K. Roy, “Low-power functionality enhanced computation architecture using spin-based devices, in *Proc. IEEE/ACM Int. Symp. Nanoscale Arch.*, pp. 129136, 2011.
- [21] D. J. Robinson “An introduction to abstract algebra”. *Walter de Gruyter*, 2003.
- [22] R. Mousavi Iraei, S. Manipatruni, D. E. Nikonov, I. A. Young and Azad Naeemi “Device and Interconnect Co-Optimization for All Spin Logic” To be submitted to *IEEE Transactions on Electron Devices*.
- [23] L. Su, Y. Zhang, J. Klein, Y. Zhang, A. Bournel, A. Fert, and W. Zhao. “Current-limiting challenges for all-spin logic devices.” *Scientific reports* 5 (2015).

- [24] C. Augustine, G. Panagopoulos, B. Behin-Aein, S. Srinivasan, A. Sarkar and K. Roy (2011, June). Low-power functionality enhanced computation architecture using spin-based devices. In Nanoscale Architectures (NANOARCH), 2011 IEEE/ACM International Symposium on (pp. 129-136). IEEE.
- [25] T. Yang, K. Kimura and Y. Otani “Giant spin-accumulation signal and pure spin-current-induced reversible magnetization switching” *Nature Phys.* 4, 851854 (2008).
- [26] A. Brataas, G. E. Bauer, and P. J. Kelly, “Non-collinear magnetoelectronics, *Phys. Rep.*, vol. 427, no. 4, pp. 157255, 2006.
- [27] S.-F. Lee, W. P. Holody, Q. Yang, P. Holody, R. Loloee, P. Schroeder, et al., “Two-channel analysis of CPP-MR data for Ag/Co and AgSn/Co multilayers, *J. Magn. Magn. Mater.*, vol. 118, nos. 12, pp. L1L5, 1993.
- [28] A. Bratas, G. E. Bauer and P. J. Kelly. ”Non-collinear magnetoelectronics”. Physics Reports, 427(4), 157-255. 2006
- [29] W. Brown, “Thermal fluctuation of fine ferromagnetic particles, *IEEE Trans. Magn.*, vol. 15, no. 5, pp. 11961208, Sep. 1979.
- [30] J. Z. Sun, “Spin-current interaction with a monodomain magnetic body: A model study, *Phys. Rev. B*, vol. 62, pp. 570578, Jul. 2000.
- [31] M. Beleggia, M. D. Graef, and Y. T. Millev, “The equivalent ellipsoid of a magnetized body, *J. Phys. D, Appl. Phys.*, vol. 39, no. 5, p. 891, 2006.
- [32] S. Rakheja, S.-C. Chang, and A. Naeemi, “Impact of dimensional scaling and size effects on spin transport in copper and aluminum interconnects, *IEEE Trans. Electron Devices*, vol. 60, no. 11, pp. 39133919, Nov. 2013.
- [33] J. Wang, H. Meng and J. P. Wang .“Programmable spintronics logic device based on a magnetic tunnel junction element.” Journal o f applied physics, 97(10), 10D509. 2005.